An exponential time upper bound for Quantum Merlin-Arthur games with unentangled provers

Martin Schwarz[1]
Dahlem Center for Complex Quantum Systems
Freie Universität Berlin, 14195 Berlin, Germany

Abstract

We prove a deterministic exponential time upper bound for Quantum Merlin-Arthur games with k unentangled provers. This is the first non-trivial upper bound of $\mathsf{QMA}(k)$ better than NEXP and can be considered an exponential improvement, unless $\mathsf{EXP} = \mathsf{NEXP}$. The key ideas of our proof are to use perturbation theory to reduce the $\mathsf{QMA}(2)$-complete SEPARABLE SPARSE HAMILTONIAN problem to a variant of the SEPARABLE LOCAL HAMILTONIAN problem with an exponentially small promise gap, and then to decide this instance using ε-net methods. Our results imply an exponential time algorithm for the PURE STATE N-REPRESENTABILITY problem in quantum chemistry, which is in $\mathsf{QMA}(2)$ but is not known to be in QMA. We also discuss the implications of our results on the BEST SEPARABLE STATE problem.

1 Introduction

Non-determinism is a fundamental concepts of theoretical computer science and led to the definition of NP, kicking of modern computational complexity theory in the 1970's [Coo71, Lev73]. Another powerful concept is interaction, where the prover/verifier interpretation of NP has been generalized to include randomness (MA, Merlin-Arthur games) [Bab85], multiple rounds of interaction (IP) [GMR85], and multiple provers (MIP) [BOGKW88]. The inherent power of these concepts became only manifest when IP and MIP were related to and characterized by well-known complextiy classes – much more powerful than NP – due to the seminal results that $\mathsf{MIP} = \mathsf{NEXP}$ [BFL91] and $\mathsf{IP} = \mathsf{PSPACE}$ [Sha92].

In the 1990's, the formal foundations of quantum computing have been laid [BV93] and analogous questions about the power of non-determinism (QMA) [Wat00, KSV02] and interaction with one or multiple provers (QIP, QMIP, and variants thereof) [Wat99, KM02] have been asked in a quantum context, and partially answered ($\mathsf{QMA} \subseteq \mathsf{PP}$, $\mathsf{QIP} = \mathsf{PSPACE}$) [MW05, JJUW10]. But in addition to that, quantum theory turned out to offer new exciting possibilities, which have no classical counterpart!

In 2001, Kobayashi, Matsumoto, and Yamakami [KMY01, KMY03] first noticed the potential computational power that might be harnessed in Quantum Merlin-Arthur games from the promise of multiple *unentangled* quantum proofs, a concept which only makes sense in the quantum setting. The resulting complexity class for k unentangled provers is called $\mathsf{QMA}(k)$, which was later shown to equal $\mathsf{QMA}(2)$ [HM13]. Liu, Christandl, and Verstraete [LCV07] showed that a natural problem in quantum chemistry, the PURE STATE N-REPRESENTABILITY problem, is in $\mathsf{QMA}(k)$ yet not obviously in QMA. Chailloux and Sattath [CS12] showed that the SEPARABLE SPARSE HAMILTONIAN problem is $\mathsf{QMA}(2)$-complete. Blier and Tapp [BT09] provided an example for the power of this class even if restricted to *tiny* proof states: they showed that NP is contained in a $\mathsf{QMA}_{\log}(2)$. In this setting, Merlin receives just two *logarithmically* sized quantum witness states relative to the input size, an exponential compression of the proof size compared to the classical case!

Aaronson et. al. [ABD+08] studied *The Power of Unentanglement* and raised the question whether the containment of NP in $\mathsf{QMA}_{\log}(2)$ might be "scaled up exponentially", such that NEXP would be contained

[1]email: martin.schwarz1@fu-berlin.de

in $\mathsf{QMA}(2)$ with *polynomially* sized quantum proofs in turn. One obstacle to reach such a conclusion is the vanishing promise gap in known reductions of NP to $\mathsf{QMA}_{\log}(2)$, whereas a *constant* gap and $O(\log(n))$-sized proofs would imply $\mathsf{QMA}(2) = \mathsf{NEXP}$. Indeed, [Per12] has shown that $\mathsf{QMA}(2)$ with exponentially small promise gap is indeed equal to NEXP. This question lead to two complementary lines of research: on the one hand, several researchers [Bei10, CF13, LGNN11] worked on improving the promise gap while maintaining a logarithmic witness size, whereas other groups started from the requirement of constant error and showed that witness sizes of $\tilde{O}(\sqrt{n})$ suffice to put NP into $\mathsf{QMA}_{\tilde{O}(\sqrt{n})}(2)$ with constant promise gap [ABD+08, HM13]. In certain restricted settings, a PSPACE upper bound for $\mathsf{QMA}(k)$ has been shown by [SW15]. Nevertheless, no non-trivial upper bound for the general class $\mathsf{QMA}(k)$ other than the trivial NEXP upper bound has been found so far.

We answer the $\mathsf{QMA}(k) \stackrel{?}{=} \mathsf{NEXP}$ question in the negative (unless $\mathsf{EXP} = \mathsf{NEXP}$) by showing:

Theorem 1. $\mathsf{QMA}(k) \subseteq \mathsf{EXP}$

Techniques The key tool we use in our proofs is the application of matrix perturbation theory to a $\mathsf{QMA}(2)$-complete SEPARABLE SPARSE HAMILTONIAN in order to reduce the locality of its globally acting sparse terms while accepting the exponential cost in operator norm.

Perturbation theory has been introduced into quantum complexity theory before by the seminal works of [KKR06, OT08, BDLT08] and we refer to these works for a detailed introduction into this technique. Its main application so far has been to reduce local Hamiltonian terms with high but constant locality to lower constant locality (e.g. 5-local to 2-local). One reason for this is the well-known fact, that these gadget constructions induce large operator norms that scale exponentially with the locality of the input terms. Thus, at most $O(\log(n))$-local terms can be reduced to constant locality while simultaneously maintaining polynomial scaling of the operator norm with the system size.

To our knowledge, this is the first work that explores the application of perturbative gadgets on *globally* acting Hamiltonian terms in the context of quantum complexity theory while accepting the exponential cost in operator norm. The resulting operators may be deemed unphysical, yet we can afford to work with them as our ultimate goal is a classical algorithm and not a physical Hamiltonian. The resulting SEPARABLE LOCAL HAMILTONIAN instances with large norm or, equivalently, small promise gap, can then be directly solved using ε-net methods [SW15] in exponential time.

Overview of the proof To show Theorem 1, we start from a generic $\mathsf{QMA}(k)$ verifier circuit. The proof then proceeds in four steps: first, we amplify the soundness and completeness bounds of the given $\mathsf{QMA}(k)$ verifier using the amplification method of Harrow and Montanaro (Theorem 8) to the levels required for the reduction to SEPARABLE SPARSE HAMILTONIAN by Chailloux and Sattath (Lemma 9). Second, we apply the Chailloux-Sattath construction yielding a SEPARABLE SPARSE HAMILTONIAN instance that contains k non-local but sparse Hamiltonian terms. Third, we apply our main technical lemma (Lemma 14) yielding a SEPARABLE LOCAL HAMILTONIAN instance with exponentially small promise gap and polynomially increased system size. Fourth, we apply the ε-net methods of Shi and Wu (6) to decide the instance in exponential time.

Structure of the paper In Section 1, we have motivated the study of $\mathsf{QMA}(k)$ and its relation to other problems, reviewed related work, presented our key results and techniques, and gave an overview of the poof of our main theorem. In Section 2, we introduce all technical definitions and earlier results that we will use in Section 3 to prove our claims. In Section 4 we discuss the implications of our results on the BEST SEPARABLE STATE problem, and finally we conclude in Section 5.

2 Preliminaries and definitions

Throughout the paper we use logarithms of base 2 and write $\tilde{O}(n) = O(n \operatorname{poly}\log(n))$. We say a pure state $|\psi\rangle$ is a *product* state, if it can be written as $|\psi\rangle = |\psi_A\rangle \otimes |\psi_B\rangle$. More generally, a mixed state (or a general operator) ρ is called *separable*, if it can be written as $\rho = \sum_i p_i \rho_i^A \otimes \rho_i^B$ with $p_i \geq 0$ and $\sum_i p_i = 1$.

Definition 2 (QMA(k))**.** *A language L is in $\mathsf{QMA}_\ell(k)_{s,c}$ if there exists a polynomial-time quantum algorithm $\mathcal{A}$ such that, for all inputs $x \in \{0,1\}^n$:*

1. **Completeness***: If $x \in L$, there exist k witnesses $|\psi_1\rangle, \ldots, |\psi_k\rangle$, each a state of ℓ qubits, such that $\mathcal{A}$ outputs* ACCEPT *with probability at least c on input $|x\rangle|\psi_1\rangle, \ldots, |\psi_k\rangle$.*
2. **Soundness***: If $x \notin L$, then $\mathcal{A}$ outputs* ACCEPT *with probability at most s on input $|x\rangle|\psi_1\rangle, \ldots, |\psi_k\rangle$, for all states $|\psi_1\rangle, \ldots, |\psi_k\rangle$.*

We use $\mathsf{QMA}(k)$ as shorthand for $\mathsf{QMA}_{\operatorname{poly}(n)}(k)_{\frac{1}{3},\frac{2}{3}}$, and QMA as shorthand for $\mathsf{QMA}(1)$. We always assume $1 \leq k \leq \operatorname{poly}(n)$. We also use the notation $\mathsf{QMA}_\ell^{\mathrm{SEP}}(k)_{s,c}$ where SEP indicates that $\mathcal{A}$'s measurement $\{M, \mathbb{1} - M\}$ is restricted to the set of separable operators $M = \sum_i \alpha_i \otimes \beta_i$ for some positive semidefinite matrices α_i, β_i.

Definition 3 (LOCAL HAMILTONIAN problem)**.** *Input: a set of hermitian matrices $H_1, \ldots, H_m$, where each matrix acts on a set of at most k out of the n qubits, and $I \succeq H_i \succeq 0$ (i.e. both H_i and $I - H_i$ are positive semidefinite), and two real number a and b such that $b - a > \operatorname{poly}(1/n)$. We define the Hamiltonian $H = \sum_{i=1}^m H_i$, where each matrix H_i is implicitly extended to the entire Hilbert space of the n qubits by tensoring identities. Output: Output* YES *if there exists a state $|\psi\rangle$ such that $\langle\psi|H|\psi\rangle \leq a$, and* NO *if for every state $|\psi\rangle$, $\langle\psi|H|\psi\rangle \geq b$. The difference $b - a$ is called the promise gap of the Hamiltonian.*

Definition 4 (SEPARABLE LOCAL HAMILTONIAN problem)**.** *The input is the same as the input for the* LOCAL HAMILTONIAN *problem together with a partition of the qubits to disjoint sets A and B. The answer is* YES *if $\exists|\psi\rangle = |\chi_A\rangle \otimes |\chi_B\rangle$ $s.t.$ $\langle\psi|H|\psi\rangle \leq a$ and the answer is* NO *if $\langle\psi|H|\psi\rangle \geq b$ for all tensor product states $|\psi\rangle = |\chi_A\rangle \otimes |\chi_B\rangle$.*

Note, that equivalent definitions can be made based on tensor products of mixed states, or separable states as witness states. We now define the SEPARABLE SPARSE HAMILTONIAN problem.

Definition 5 (SEPARABLE SPARSE HAMILTONIAN problem)**.** *An operator A over n qubits is row-sparse if each row in A has at most $\operatorname{poly}(n)$ non-zero entries, and there exists an efficient classical algorithm that, given i, outputs a list $(j, A_{i,j})$ running over all non zero elements of $A_{i,j}$. The* SEPARABLE SPARSE HAMILTONIAN *problem is the same as* SEPARABLE LOCAL HAMILTONIAN *except each term H_i in the input Hamiltonian is row-sparse instead of k-local.*

We will use the following theorem of Shi and Wu [SW15].

Corollary 6 ([SW15, Problem 3, Corollary 6])**.** *Take the expression $Q = \sum_{i=1}^r H_i$ of any l-local Hamiltonian over $A_1 \otimes \cdots \otimes A_k$ (each A_i is of dimension $d = 2^n$) such that $\|H_i\|_{\mathrm{op}} \leq w$ for each i as input. Assuming $k, l = O(1)$, the quantity*

$$OptSep(Q) = \min \langle Q, \rho \rangle \text{ subject to } X \in SepD(A_1 \otimes \cdots \otimes A_k) \tag{1}$$

where $SepD(A_1 \otimes \cdots \otimes A_k)$ is the set of separable density operators over the space $A_1 \otimes \cdots \otimes A_k$, can be approximated to precision δ in

$$\mathsf{DTIME}(\exp(O(\log^{O(1)}(d)(\log\log(d) + \log(w/\delta)))) \times \operatorname{poly}(d, w, 1/\delta)), \tag{2}$$

which is quasi-polynomial in $d, w, 1/\delta$. If n is considered as the input size and $w/\delta = O(\operatorname{poly}(n))$, then OptSep(Q) can be approximated to precision δ in PSPACE.

Harrow and Montanaro [HM13] prove the following results:

Lemma 7 ([HM13, Lemma 6]). *For any $m, k, 0 \leq s < c \leq 1$,*

$$\mathsf{QMA}_m(k)_{s,c} \subseteq \mathsf{QMA}^{\mathrm{SEP}}_{km}(2)_{s',c'}$$

where $c' = \frac{1+c}{2}$ and $s' = 1 - \frac{(1-s)^2}{100}$.

Theorem 8 ([HM13, Theorem 9]).

1. *If $s \leq 1 - 1/q(n), k = \mathrm{poly}(n)$ and $p(n), q(n)$ be arbitrary polynomials, then*[2]

$$\mathsf{QMA}_\ell(k)_{s,1} = \mathsf{QMA}^{\mathrm{SEP}}_{O(k\ell p(n)q^2(n))}(2)_{exp(-p(n)),1}. \tag{3}$$

2. *If $c - s \geq 1/q(n), c < 1, k = \mathrm{poly}(n)$ and $p(n), q(n)$ be arbitrary polynomials, then*

$$\mathsf{QMA}_\ell(k)_{s,c} = \mathsf{QMA}^{\mathrm{SEP}}_{\tilde{O}(k\ell p^2(n)q^2(n))}(2)_{exp(-p(n)),1-exp(-p(n))}. \tag{4}$$

Lemma 9 (SEPARABLE SPARSE HAMILTONIAN is $\mathsf{QMA}(2)$-hard [CS12]). *Let $U = U_T U_{T-1} \dots U_0$ be the verifier circuit of a language $L \in \mathsf{QMA}^{SEP}_\ell(2)_{\frac{1}{T+1}, 1 - \frac{C}{512(T+1)^4}}$ for some constant C with input x, proof size ℓ, and α ancillas. W.l.o.g., assume that U has been produced by Lemma 7 or Theorem 8. Then there exists a separable sparse Hamiltonian H_{SSH} that is a sum of $O(T)$ sparse terms acting on at most $2\ell + O(T) + \alpha$ qubits, such that*

1. *if $x \in L$ then there exists a $|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle$, such that $\langle\psi|H|\psi\rangle \leq \frac{C}{512(T+1)^5}$.*
2. *if $x \notin L$ then for all $|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle$, $\langle\psi|H|\psi\rangle \geq \frac{C}{256(T+1)^5}$*

Theorem 10 (Eigenvalue-Approximating Gadget Theorem [OT08, Theorem 7]). *Let $||V|| \leq \Delta/2$ where Δ is the spectral gap of H and $\lambda(H) = 0$. Let $\tilde{H}|_{<\Delta/2}$ be the restriction of $\tilde{H} = H + V$ to the space of eigenstates with eigenvalues less than $\Delta/2$. Let there be an effective Hamiltonian H_{eff} with $\mathrm{Spec}(H_{\mathrm{eff}}) \subseteq [a, b]$. If the self-energy $\Sigma_-(z)$ for all $z \in [a - \varepsilon, b + \varepsilon]$ where $a < b < \Delta/2 - \varepsilon$ for some $\varepsilon > 0$, has the property that*

$$||\Sigma_-(z) - H_{\mathrm{eff}}|| \leq \varepsilon, \tag{5}$$

then each eigenvalue $\tilde{\lambda}_j$ of $\tilde{H}|_{<\Delta/2}$ is ε-close to the jth eigenvalue of H_{eff}. In particular

$$|\lambda(H_{\mathrm{eff}}) - \lambda(\tilde{H})| \leq \varepsilon. \tag{6}$$

Theorem 11 (Norm-Approximating Gadget Theorem [OT08, Theorem A.1]). *Given is a Hamiltonian H such that no eigenvalues of H lie between $\lambda_- = \lambda_* - \Delta/2$ and $\lambda_+ = \lambda_* + \Delta/2$. Let $\tilde{H} = H + V$ where $||V|| \leq \Delta/2$. Let there be an effective Hamiltonian H_{eff} with $\mathrm{Spec}(H_{\mathrm{eff}}) \subseteq [a, b]$, $a < b$. We assume that $H_{\mathrm{eff}} = \Pi_- H_{\mathrm{eff}} \Pi_-$. Let D_r be a disk of radius r in the complex plane centered around $z_0 = \frac{b+a}{2}$. Let r be such that $b + \varepsilon < z_0 + r < \lambda_*$. Let $w_{\mathrm{eff}} = \frac{b-a}{2}$. Assume that for all $z \in D_r$ we have*

$$\|\Sigma_-(z) - H_{\mathrm{eff}}\| \leq \varepsilon. \tag{7}$$

Then

$$\|\tilde{H}_{<\lambda_*} - H_{\mathrm{eff}}\| \leq \frac{3(||H_{\mathrm{eff}}|| + \varepsilon)\|V\|}{\lambda_+ - ||H_{\mathrm{eff}}|| - \varepsilon} + \frac{r(r + z_0)\varepsilon}{(r - w_{\mathrm{eff}})(r - w_{\mathrm{eff}} - \varepsilon)}. \tag{8}$$

[2] We have explicitly included the asymptotic scaling of the proof sizes, which are implicit in the proof of the theorem in [HM13] but were not included in the original statement of the theorem.

Lemma 12 (Norm-approximating Parallel Subdivision Gadget [OT08])**.** *Let $H_{\text{target}} = H_{\text{else}} + \sum_{i=1}^{k} A_i \otimes B_i$ be an ℓ-local Hamiltonian, where A_i and B_i are k-many pairs of $(\ell/2)$-local terms and H_{else} contains all terms that shall not be generated perturbatively. Let $H_{\text{eff}} = H_{\text{target}} \otimes |0..0\rangle\langle 0..0|$ be an effective Hamiltonian acting on a larger system extended by $O(k)$ ancillas, i.e. H_{target} acting on the subspace where the ancillas are in their ground state. Let $\Delta = \text{poly}(n,k)/\varepsilon^2$ for a sufficiently large $\text{poly}(n,k)$. Then for any ε there exists a $(\lceil \ell/2 \rceil + 1)$-local Hamiltonian $\tilde{H}$ with $\|\tilde{H}\| = \Delta$ such that*

$$\left\| \tilde{H}|_{<\Delta/2} - H \otimes |0...0\rangle\langle 0...0| \right\| \leq \varepsilon \tag{9}$$

where $\tilde{H}|_{<\Delta/2}$ indicates the restriction of $\tilde{H}$ to the space of eigenvectors with eigenvalues less than $\Delta/2$.

Proof sketch. Lemma 12 is implicit in [OT08, Appendix A]: Consider the parallel subdivision gadget construction of [OT08, Section 3.1]. In the construction, Theorem 10 is applied to construct an operator $\tilde{H}$ such that each eigenvalue $\tilde{\lambda}_j$ of $\tilde{H}|_{<\Delta/2}$ is ε-close to the jth eigenvalue of H_{eff}. It is straightforward to check that the same assumptions (up to a larger polynomial in the choice of Δ) suffice to apply Theorem 11 instead of Theorem 10 to the constructed gadget Hamiltonian, yielding the norm approximation $\left\| \tilde{H}|_{<\Delta/2} - H \otimes |0...0\rangle\langle 0...0| \right\| \leq \varepsilon$ instead of an eigenvalue approximation. □

3 Proof

We will now proceed to prove Theorem 1 by showing the following slightly more general result.

Theorem 13 (main result)**.** $\mathsf{QMA}_{\ell}(k)_{s,c}$ *with $c - s > 1/q$, $q = q(n)$, is decidable in*

$$\mathsf{DTIME}(\exp(O(\text{poly}(k, \ell, q, \alpha, T, \log(n))))) \subseteq \mathsf{EXP}, \tag{10}$$

where T is a bound on the size of the $\mathsf{QMA}(k)$ verifier circuit and α a bound on the number of ancillas used.

We note that our upper bound is consistent with previously known hardness results, such as $\mathsf{NEXP} \subseteq \mathsf{QMA}(2)_{s,c}$ with $c - s \geq 2^{-O(\text{poly}(n))}$ [Per12] and $\mathsf{NP} \subseteq \mathsf{QMA}_{\log}(2)_{s,c}$ with $1/(c-s) \leq O(\text{poly}(n))$ [BT09, BT12], as well as the general lower bounds of [HM13, AIM14] relative to the Exponential-Time Hypothesis (ETH) of [IP99].

Proof. The proof proceeds in four steps: first, we amplify the soundness and completeness bounds of the given $\mathsf{QMA}(k)$ verifier using Theorem 8 to the levels required by Lemma 9. Second, we apply Lemma 9 yielding a SEPARABLE SPARSE HAMILTONIAN instance. Third, this serves as input to Lemma 14 which yields a SEPARABLE LOCAL HAMILTONIAN instance with exponentially small promise gap. Fourth, we apply 6 and decide the instance in EXP. Let us now discuss these steps in detail.

Step one ($\mathsf{QMA}(k)$ amplification). We apply Theorem 8 to the $\mathsf{QMA}_{\ell}(k)_{s,c}$ verifier circuit, yielding a $\mathsf{QMA}^{\text{SEP}}_{\ell'}(2)_{e^{-p(n)},1-e^{-p(n)}}$ verifier circuit, where the proof size has been expanded to $\ell' = \tilde{O}(k\ell p^2(n) q^2(n))$. Choosing $p(n) = 10\log(T) + D$ for a sufficiently large constant D satisfies the bounds $s > \frac{1}{T+1}$ and $c < \frac{1}{512(T+1)^4}$ required by Lemma 9. Thus we have $\ell' = \tilde{O}(k\ell q^2(n)\log^2(T))$

Step two (reduction to SEPARABLE SPARSE HAMILTONIAN). We apply Lemma 9 to the $\mathsf{QMA}^{\text{SEP}}_{\ell'}(2)_{s,c}$ instance of step one, with $c = \frac{1}{T+1}$, $s = \frac{1}{512(T+1)^4}$, $\ell' = \tilde{O}(k\ell \log^2(T) q^2(n))$, yielding a separable local Hamiltonian H_{SSH} with energy thresholds $a \leq \frac{C}{512(T+1)^5}$ and $b \geq \frac{C}{256(T+1)^5}$ in the YES and NO cases, respectively, where H_{SSH} acts on $w = O(2\ell' + \alpha + T) = \tilde{O}(k\ell \log^2(T) q^2(n) + \alpha + T)$ qubits, where α is the number of ancilla bits used by the original verifier.

Step three (reduction to SEPARABLE LOCAL HAMILTONIAN). We apply our main technical Lemma 14 to H_{SSH} produced in step two. This yields a SEPARABLE LOCAL HAMILTONIAN instance H_{SLH} with ground energies $a \leq \frac{5C}{2048(T+1)^5}\, 2^{-O(\text{poly}(\ell')\log(nkT))}$ and $b \geq \frac{7C}{2048(T+1)^5}\, 2^{-O(\text{poly}(\ell')\log(nkT))}$ in the YES and NO cases, respectively, where H_{SLH} acts on $w' = O(2k\ell' + \alpha + T) = \tilde{O}(k^2\ell\log^2(T)q^2(n) + \alpha + T)$ qubits.

Step four (enumeration of ε-net). Finally, we apply 6 to H_{SLH} to approximate the ground energy of H_{SLH} over the set of separable states to precision

$$\delta = \frac{C}{2048(T+1)^5}\, 2^{-O(\text{poly}(\ell')\log(nkT))} \tag{11}$$

$$\approx 2^{-O(\text{poly}(\ell')\log(nkT))} \tag{12}$$

which suffices to decide the H_{SLH} instance. Since $d = 2^{w'}$, this requires at most

$$\mathsf{DTIME}(\exp(O(\log^{O(1)}(2^{w'})(\log\log(2^{w'}) + \log(1/\delta)))) \times \text{poly}(2^{w'}, 1/\delta)) \tag{13}$$

Expanding all parameters we find an upper bound of

$$\mathsf{DTIME}(\exp(O(\text{poly}(k, \ell, q, \alpha, T, \log(n))))) \tag{14}$$

Clearly, for $\mathsf{QMA}(k)$ with $k, \ell, q, \alpha, T \in O(\text{poly}(n))$ we have

$$\mathsf{QMA}(k) \subseteq \mathsf{DTIME}(\exp(O(\text{poly}(n)))) \subseteq \mathsf{EXP}. \tag{15}$$

Note, that the condition in 6 implying containment in PSPACE is not satisfied in our setting. □

It remains to prove our main technical lemma.

Lemma 14 (SEPARABLE LOCAL HAMILTONIAN with exponentially small promise gap is $\mathsf{QMA}(2)$-hard)**.** *Let H_{SSH} be a* SEPARABLE SPARSE HAMILTONIAN *instance produced by Lemma 9. Then there exists a* SEPARABLE LOCAL HAMILTONIAN *instance H_{SLH} such that*

1. *if $x \in L$ then there exists a $|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle$, such that $\langle\psi|H|\psi\rangle \leq \frac{5C}{2048(T+1)^5}\, 2^{-O(\text{poly}(\ell)\log(nkT))}$*
2. *if $x \notin L$ then for all $|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle$, $\langle\psi|H|\psi\rangle \geq \frac{7C}{2048(T+1)^5}\, 2^{-O(\text{poly}(\ell)\log(nkT))}$*

H_{SLH} acts on an enlarged system of $O(2k\ell + T + \alpha)$ qubits.

Proof. Note that H_{SSH} is a Feynman-Kitaev Hamiltonian [KSV02, section 14.4.1] of the standard form

$$H_{\text{SSH}} = H_{\text{in}} + H_{\text{prop}} + H_{\text{out}} + H_{\text{clock}} \tag{16}$$

where $H_{\text{prop}} = \sum_{t=1}^{T} H_t$. Since by the assumptions of Lemma 9 the verifier circuit encoded in H_{SSH} has been brought into a standard form by the Harrow-Montanaro construction (i.e. Lemma 7 or Theorem 8), we know that H_{SSH} contains exactly k non-local terms (one for each prover) of the form

$$H_t = -\frac{1}{2}|t\rangle\langle t-1| \otimes U_t - \frac{1}{2}|t-1\rangle\langle t| \otimes U_t^\dagger + \frac{1}{2}(|t\rangle\langle t| + |t+1\rangle\langle t+1|) \otimes \mathbb{1} \tag{17}$$

These encode a simultaneous controlled-swap operation $U_t = \text{CSWAP}$ on 2ℓ qubits. All other terms are 5-local. As noticed by Chailloux and Sattath [CS12], it's necessary to perform this controlled-swap operation

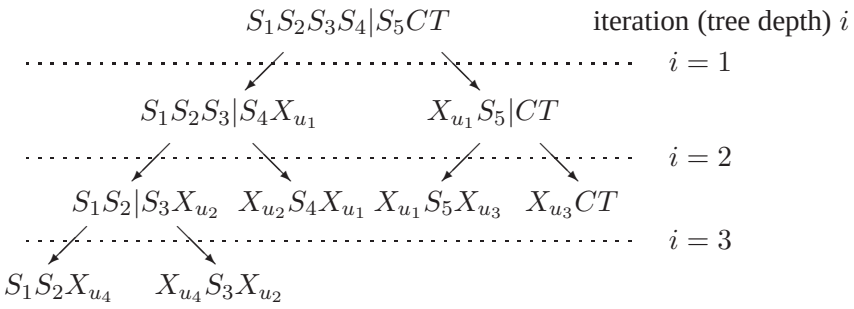

Figure 1: [CBBK15] Reduction tree diagram for parallel subdivision gadget acting on a H_{CSWAP} term as defined in eq. (18) for the case $\ell = 5$. In our case, each S_i is a two-qubit SWAP operator acting on qubits $(i, \ell + i)$, C is the control operator, whereas T is the time propagation operator. The vertical lines $|$ show where the subdivisions are made at each iteration to each term. The X_{u_i} terms indicate the coupling to mediator qubit u_i introduced in this step. Clearly, the number of mediator qubits scales as $O(\ell)$ and $O(\log(\ell))$ iterations suffice to arrive at $O(1)$-local terms.

"in one time step" in the history state that constitutes the ground state of the Feynman-Kitaev Hamiltonian. This is required in order to ensure the separability of the ground state of H_{SSH} for satisfiable instances.

From this starting point, we show the lemma by reducing SEPARABLE SPARSE HAMILTONIAN to SEPARABLE LOCAL HAMILTONIAN with exponentially small promise gap. Our approach to deal with the non-local terms is to approximate them by local ones *perturbatively* using the parallel subdivision gadget of [OT08] as summarized in Lemma 12. Let us extract one representative non-local term H_{CSWAP} from H_t. The lemma is later applied to all k of these terms in parallel. Note that the CSWAP operation across 2ℓ qubits exhibits a natural tensor product structure which is a crucial prerequisite to apply Lemma 12 iteratively. Using $\mathrm{CSWAP} = \mathrm{CSWAP}^\dagger$ we write

$$H_{\mathrm{CSWAP}} = \overbrace{\mathrm{SWAP}_{1,\ell+1} \otimes \mathrm{SWAP}_{2,\ell+2} \otimes \cdots \otimes \mathrm{SWAP}_{\ell,2\ell}}^{\text{swap terms } S_i} \otimes \overbrace{|1\rangle\langle 1|}^{\text{control } C} \otimes \overbrace{(|t\rangle\langle t+1| + |t+1\rangle\langle t|)}^{\text{time propagation } T} \tag{18}$$

We apply Lemma 12 to H_{SSH} iteratively $O(\log(\ell))$ times in order to break down H_{CSWAP} into ultimately $O(1)$-local terms along its natural tensor product structure as illustrated in Figure 1. For one application of Lemma 12 a choice of

$$\Delta_1 = \frac{\mathrm{poly}(n,k)}{\varepsilon^2} \tag{19}$$

suffices. Iterating the gadget increases the required interaction strength by a polynomial factor [BDLT08].

$$\Delta_{i+1} = \frac{\mathrm{poly}(n,k)}{\varepsilon^2}\Delta_i^3 \tag{20}$$

Thus, after $O(\log(\ell))$ iterations, we find

$$\Delta \leq \left(\frac{\mathrm{poly}(n,k)}{\varepsilon^2}\right)^{3^{\log(\ell)}} \leq \left(\frac{\mathrm{poly}(n,k)}{\varepsilon^2}\right)^{\mathrm{poly}(\ell)} \leq 2^{O(\mathrm{poly}(\ell)\log(nk/\varepsilon))} \tag{21}$$

It suffice to choose $\varepsilon = \frac{C}{2048(T+1)^5\log(\ell)}$ to resolve the promise gap of H_{SSH} (see Lemma 9) considering $O(\log(\ell))$ parallel iterations per term, each incurring error ε. The result of the iterated gadget construction is a 3-local operator $\tilde{H}_{\mathrm{SSH}}$ with

$$\left\|\tilde{H}_{\mathrm{SSH}}\right\| \leq \Delta \leq 2^{O(\mathrm{poly}(\ell)\log(nkT))} \tag{22}$$

such that

$$\left\| \tilde{H}_{\text{SSH}} - H_{\text{SSH}} \otimes |0\cdots 0\rangle\langle 0\cdots 0| \right\| \leq \frac{C}{2048(T+1)^5} \tag{23}$$

Figure 1 illustrates how $O(\ell)$ mediator qubits are introduced per term. Since there are k non-local terms being reduced in parallel, a total of $O(k\ell)$ mediator qubits are added. Let us finally define the desired SEPARABLE LOCAL HAMILTONIAN instance as the normalized version of $\tilde{H}_{\text{SSH}}$ as

$$H_{\text{SLH}} = \Delta^{-1}\tilde{H}_{\text{SSH}} \tag{24}$$

To finish the proof of the lemma, it remains to show that the claimed *completeness* and *soundness* bounds are satisfied by H_{SLH}. Let us first verify completeness and soundness of $\tilde{H}_{\text{SSH}}$ explicitly which will in turn imply the exponentially rescaled bounds for H_{SLH}.

For the completeness bound we show the that the separable witness state implied by Lemma 9, i.e. $|\psi\rangle = |\psi_A\rangle \otimes |\psi_B\rangle$ with energy $\langle\psi|H_{\text{SSH}}|\psi\rangle \leq a$, implies that the separable state $|\psi\rangle|0\cdots 0\rangle$ of the extended system satisfies $\langle\psi|\langle 0\cdots 0|\ \tilde{H}_{\text{SSH}}\ |\psi\rangle|0\cdots 0\rangle \leq \frac{5C}{2048(T+1)^5}$. Omitting the explicit tensoring of $|0\cdots 0\rangle$ ancillas for readability, we have

$$\begin{aligned}
\langle\psi|\tilde{H}_{\text{SSH}}|\psi\rangle &= \langle\psi|H_{\text{SSH}}|\psi\rangle + \langle\psi|(\tilde{H}_{\text{SSH}} - H_{\text{SSH}})|\psi\rangle && (25)\\
&\leq \frac{C}{512(T+1)^5} + \left\|\tilde{H}_{\text{SSH}} - H_{\text{SSH}}\right\| && (26)\\
&\leq \frac{4C}{2048(T+1)^5} + \frac{C}{2048(T+1)^5} && (27)\\
&= \frac{5C}{2048(T+1)^5} && (28)
\end{aligned}$$

where the first inequality follows from the assumptions of Lemma 9 and uses basic properties of the spectral norm, while the second inequality follows from eq. (23). Note that essentially the same separable state $|\psi\rangle = |\psi_A\rangle \otimes |\psi_B\rangle$ (once extended to the larger space with $|0\cdots 0\rangle$ ancillas) satisfies the respective completeness bounds in both, the perturbed as well as the unperturbed setting.

Similarly, for the soundness bound, we know that for *all* states $|\psi\rangle$

$$\begin{aligned}
\langle\psi|H_{\text{SSH}}|\psi\rangle &= \langle\psi|\tilde{H}_{\text{SSH}}|\psi\rangle + \langle\psi|(H_{\text{SSH}} - \tilde{H}_{\text{SSH}})|\psi\rangle && (29)\\
&\leq \langle\psi|\tilde{H}_{\text{SSH}}|\psi\rangle + \left\|H_{\text{SSH}} - \tilde{H}_{\text{SSH}}\right\| && (30)\\
&\leq \langle\psi|\tilde{H}_{\text{SSH}}|\psi\rangle + \frac{C}{2048(T+1)^5} && (31)
\end{aligned}$$

where eq. (30) follows from basic properties of the spectral norm, and eq. (31) follows from eq. (23). Furthermore, since for all separable $|\psi\rangle = |\psi_A\rangle \otimes |\psi_B\rangle$ we have $\langle\psi|H_{\text{SSH}}|\psi\rangle \geq \frac{C}{256(T+1)^5}$ by the assumptions of Lemma 9, it follows that

$$\langle\psi|\tilde{H}_{\text{SSH}}|\psi\rangle \geq \frac{7C}{2048(T+1)^5} \tag{32}$$

Dividing eq. (28) and eq. (32) by Δ already yields the lemma. Moreover, for H_{SLH} these bounds imply a promise gap of

$$\gamma = \frac{1}{\Delta}\frac{7C - 5C}{2048(T+1)^5} = \frac{2C}{2048(T+1)^5}\ 2^{-O(\text{poly}(\ell)\log(nkT))} \tag{33}$$

Clearly, the inverse exponential scaling of γ in ℓ dominates the scaling in all other parameters. Since in general QMA(2) instances $\ell, k, T \in O(\text{poly}(n))$, the promise gap of H_{SLH} scales with $2^{-O(\text{poly}(n))}$. □

4 On the BEST SEPARABLE STATE problem

In this section, we review the implications of our results on the related BEST SEPARABLE STATE problem. The complexity of QMA(2) stems from essentially two sources: the search for a witness state over the set of *separable* states and the fact the verifier is a quantum computer. Removing the second aspect, one is lead to following related problem, which is often discussed in the context of QMA(2).

Problem 1 (BEST SEPARABLE STATE BSS_ε)**.** *Given as input an Hermitian matrix $A \in \mathbb{C}^{d^2\times d^2}$, with eigenvalues in $[0,1]$, and let*

$$\lambda_{sep}(A) := \max_{v,w\in\mathbb{C}^d:\|v\|=\|w\|=1} (v^\dagger \otimes w^\dagger) A (v\otimes w). \tag{34}$$

Compute $\tilde{\lambda}_{sep}(A)$, such that $|\lambda_{sep}(A) - \tilde{\lambda}_{sep}(A)| \leq \varepsilon$. (Here ε is assumed to be an arbitrarily small constant if not specified otherwise.)

This problem has been related to 18 approximately equivalent problems by Harrow and Montanaro [HM13] emphasizing the importance of understanding the complexity of BSS_ε and its relation to QMA(2). For example, the BEST SEPARABLE STATE problem is equivalent to approximating the *injective tensor norm* of a 3-index tensor [DF92], a generic problem arising in several contexts (e.g. in [BBH+12] in relation to the Unique Games Conjecture [Kho02]), variants of the PLANTED CLIQUE problem, and – unsurprisingly – various problems in quantum information theory.

Note, that BSS_ε is clearly a generalization of $\mathsf{QMA}(2)_{s,c}$ as we have removed assumptions about the input. Indeed, by choosing $d = 2^{\mathrm{poly}(n)}$ and $\varepsilon = 1/\mathrm{poly}(n)$, QMA(2) can be reduced to an exponentially large instance of BSS_ε in exponential time in n: It suffices to compute A classically by multiplying the matrices defining the verifier circuit V and choosing $\varepsilon = \frac{c-s}{2} = 1/\mathrm{poly}(n)$.

Hardness of BSS_ε What is known about the hardness of the problem? $\mathrm{BSS}_{1/\mathrm{poly}(d)}$ is already known to be NP-hard [Gur03, Ioa07, Gha10] and is closely related to the problem of entanglement detection in quantum states. A long-standing open question is whether BSS_ε remains NP-hard in the regime of constant ε. [HM13, AIM14] give a lower bound of $d^{\Omega(\log(d))}$ relative to the Exponential Time Hypothesis [IP99]. For the special case that $\|A\|_F^2 = O(\mathrm{poly}\log(d))$ or assuming that $\{A, \mathbb{1} - A\}$ is an LOCC measurement (which allows local operations and only classical communication across the subsystem boundary), quasi-polynomial time algorithms are known [BCY11, SW15].

Impact of our results Do our perturbative methods yield a quasi-polynomial time algorithm for the general case? Interestingly, this does *not* seem to be the case. Let us briefly discuss informally, why two of the most natural approaches fail to yield a better upper bound.

One natural approach is to reduce BSS_ε to an instance of $\mathsf{QMA}_{\log}(2)$ and then apply Theorem 13. If after the reduction all parameters in the application of Theorem 13 turned out to be $O(\mathrm{poly}\log(d))$ this would result in a quasi-polynomial time upper bound for BSS_ε in terms of d. Such a reduction appears to introduce insourmountable overhead, though: Since we lack a $O(\log(d))$-sized circuit decomposition of A, we can only recover a quantum circuit very generically by first diagonalizing $A = UDU^\dagger$ in the eigenbasis, and then decomposing the unitary U over some gate set using the Solovay-Kitaev algorithm [NC11]. Assuming d a power of 2, $\ell = \log(d)$, and an approximation error of ε, this will yield a quantum circuit of size

$$T = O(4\ell^2 4^{2\ell}\log^c(4\ell^2 4^{2\ell}/\varepsilon)) = O(\mathrm{poly}(d, \log(1/\varepsilon)) \tag{35}$$

acting on 2ℓ qubits. Thus, even though $k, q, \ell, \alpha, \log(n) \leq O(\log(d))$ in this case, Theorem 13 only yields a run-time bound exponential in d due to T scaling polynomial in d.[3]

[3] Note, that this rough bound does not even include the number of gates required for implementing D yet.

Another natural approach is to consider $H = \mathbb{1} - A$ as a (global) Hamiltonian with the goal to apply the perturbative gadgets immediately. Clearly, $0 \leq H \leq \mathbb{1}$ and approximating $1 - \lambda_{sep}(A)$ to additive error ε is equivalent to approximating the ground energy of H over the set of separable states. Since the perturbative gadgets require a tensor product structure in the Hamiltonian terms, decompose H over the Pauli product operator basis, i.e. $H = \sum c_{i_1,\ldots,i_n} \sigma_{i_1} \otimes \cdots \otimes \sigma_{i_\ell}$, where $\ell = \log(d)$ respecting the tensor product structure of $v \otimes w$. Without further assumptions, the sum consists of $4^{\log(d^2)} = d^4$ terms. Using Lemma 12 iteratively $O(\log(\ell))$ times we can now break down each of the d^4 global terms in H into $O(n)$ 3-local terms at the cost of increased operator norm. Furthermore, each non-local term induces $O(\log(d))$ mediator qubits. Thus, after the reduction there are $O(d^4 \log(d))$ terms in a transformed local Hamiltonian $\tilde{H}$ acting on an enlarged system of $O(d^4 \log(d))$ qubits or dimension $d' = 2^{O(d^4 \log(d))}$. To approximate the operator in norm within ε it suffices to choose (cf. eq. (21))

$$\Delta = 2^{O(\mathrm{poly}(\ell) \log(nk/\varepsilon))} = 2^{O(\mathrm{poly}(\log(d)) \log(d^4 \log(d)/\varepsilon))} \leq 2^{O(\mathrm{poly}(\log(d), \log(1/\varepsilon)))} \tag{36}$$

in Lemma 12. This yields the operator $\tilde{H}$ with $\left\| H - \tilde{H} \right\| \leq \varepsilon$ on the low-energy subspace of interest. Then we apply 6, which allows us to approximate $1 - \lambda_{sep}(A)$ to precision $O(\varepsilon)$ in

$$\mathsf{DTIME}(\exp(O(\log^{O(1)}(d')(\log\log(d') + \log(\Delta/\varepsilon)))) \times \mathrm{poly}(d', \Delta, 1/\varepsilon)), \tag{37}$$

which simplifies to

$$\mathsf{DTIME}(\exp(O(\mathrm{poly}(d, \log(1/\varepsilon))) \tag{38}$$

In summary, we find that the lack of further structural information about A, such as a short circuit decomposition or a short Pauli decomposition (both in terms of $O(\mathrm{poly}\log(d))$), is a significant obstacle for solving BSS_ε in quasi-polynomial time using our methods. Only in the special case of $\mathsf{QMA}(2)$, where such information is available, our method is able to effectively exploit it and yields a quasi-polynomial time upper bound (in terms of d), but not in the general case of BSS_ε. This is consistent with earlier quasi-polynomial time algorithms for BSS_ε [BCY11, SW15] which require a bound on $\|A\|_F^2 = O(\mathrm{poly}\log(d))$ as well. Thus, we conclude that some logarithmic bound on a structural measure of A appears to be necessary to achieve a sub-exponential time bound for the problem.

5 Conclusion

We have shown the first non-trivial upper bound on $\mathsf{QMA}(k)$. In fact, we have shown how to solve the class in deterministic exponential time and ruled out its equivalence with NEXP, unless $\mathsf{NEXP} = \mathsf{EXP}$. Our results imply an exponential time algorithm for the PURE STATE N-REPRESENTABILITY problem in quantum chemistry, which is in $\mathsf{QMA}(2)$ but is not known to be in QMA. Furthermore, we have discussed the implications of our results on the BSS_ε problem and explained why no quasi-polynomial time algorithm for BSS_ε follows. Rather, we found that the quantum circuit structure present in $\mathsf{QMA}(k)$ but missing in BSS_ε is necessary to apply our techniques effectively. In this paper, we were mainly concerned with proving an exponential time upper bound for $\mathsf{QMA}(k)$ and leave the explicit determination of the polynomials in our upper bounds open for future work.

6 Acknowledgments

MS gratefully acknowledges the suggestion of Xiaodi Wu, during an informal presentation of an earlier, more lengthy version of the proof, to show step four of Theorem 13 by applying his previously published theorem [SW15, Corollary 6]. This helped shorten the paper and improved the presentation of the result. MS also acknowledges helpful discussions with Scott Aaronson, Yudong Cao, Aram Harrow, Daniel Nagaj, and Thomas Vidick as well as financial support by the Alexander-von-Humboldt Foundation. MS also thanks the EU projects RAQUEL, SIQS, AQuS for supporting research at the Dahlem Center for Complex Quantum Systems.

References

[ABD+08] Scott Aaronson, Salman Beigi, Andrew Drucker, Bill Fefferman, and Peter Shor. The power of unentanglement. In *Computational Complexity, 2008. CCC'08. 23rd Annual IEEE Conference on*, pages 223–236. IEEE, 2008.

[AIM14] Scott Aaronson, Russell Impagliazzo, and Dana Moshkovitz. AM with multiple merlins. In *Computational Complexity (CCC), 2014 IEEE 29th Conference on*, pages 44–55. IEEE, 2014.

[Bab85] L Babai. Trading group theory for randomness. In *Proceedings of the Seventeenth Annual ACM Symposium on Theory of Computing*, STOC '85, pages 421–429, New York, NY, USA, 1985. ACM.

[BBH+12] Boaz Barak, Fernando GSL Brandao, Aram W Harrow, Jonathan Kelner, David Steurer, and Yuan Zhou. Hypercontractivity, sum-of-squares proofs, and their applications. In *Proceedings of the forty-fourth annual ACM Symposium on Theory of Computing*, pages 307–326. ACM, 2012.

[BCY11] Fernando GSL Brandão, Matthias Christandl, and Jon Yard. A quasipolynomial-time algorithm for the quantum separability problem. In *Proceedings of the forty-third annual ACM symposium on Theory of computing*, pages 343–352. ACM, 2011.

[BDLT08] Sergey Bravyi, David P DiVincenzo, Daniel Loss, and Barbara M Terhal. Quantum simulation of many-body hamiltonians using perturbation theory with bounded-strength interactions. *Physical review letters*, 101(7):070503, 2008.

[Bei10] Salman Beigi. NP vs. QMA log (2). *Quantum Information & Computation*, 10(1):141–151, 2010.

[BFL91] László Babai, Lance Fortnow, and Carsten Lund. Non-deterministic exponential time has two-prover interactive protocols. *Computational complexity*, 1(1):3–40, 1991.

[BOGKW88] Michael Ben-Or, Shafi Goldwasser, Joe Kilian, and Avi Wigderson. Multi-prover interactive proofs: How to remove intractability assumptions. In *Proceedings of the twentieth annual ACM symposium on Theory of computing*, pages 113–131. ACM, 1988.

[BT09] Hugue Blier and Alain Tapp. All languages in NP have very short quantum proofs. In *Quantum, Nano and Micro Technologies, 2009. ICQNM'09. Third International Conference on*, pages 34–37. IEEE, 2009.

[BT12] Hugue Blier and Alain Tapp. A quantum characterization of NP. *computational complexity*, 21(3):499–510, 2012.

[BV93] E. Bernstein and U. Vazirani. Quantum complexity theory. In *Proceedings of the twenty-fifth annual ACM symposium on Theory of computing*, page 20. ACM, 1993.

[CBBK15] Yudong Cao, Ryan Babbush, Jacob Biamonte, and Sabre Kais. Hamiltonian gadgets with reduced resource requirements. *Physical Review A*, 91(1):012315, 2015.

[CF13] Alessandro Chiesa and Michael A Forbes. Improved soundness for qma with multiple provers. *Chicago Journal of Theoretical Computer Science*, 1:1–23, 2013.

[Coo71] Stephen A. Cook. The complexity of theorem-proving procedures. In *Proceedings of the Third Annual ACM Symposium on Theory of Computing*, STOC '71, pages 151–158, New York, NY, USA, 1971. ACM.

[CS12] Andre Chailloux and Or Sattath. The complexity of the separable hamiltonian problem. In *Proceedings of the 2012 IEEE Conference on Computational Complexity (CCC)*, pages 32–41. IEEE Computer Society, 2012.

[DF92] Andreas Defant and Klaus Floret. *Tensor norms and operator ideals*, volume 176. Elsevier, 1992.

[Gha10] Sevag Gharibian. Strong NP-hardness of the Quantum Separability Problem. *Quantum Information and Computation*, 10, 2010.

[GMR85] Shafi Goldwasser, Silvio Micali, and Charles Rackoff. The knowledge complexity of interactive proof-systems. In *Proceedings of the seventeenth annual ACM symposium on Theory of computing*, pages 291–304. ACM, 1985.

[Gur03] Leonid Gurvits. Classical deterministic complexity of edmonds' problem and quantum entanglement. In *Proceedings of the thirty-fifth annual ACM symposium on Theory of computing*, pages 10–19. ACM, 2003.

[HM13] Aram W Harrow and Ashley Montanaro. Testing product states, quantum Merlin-Arthur games and tensor optimization. *Journal of the ACM (JACM)*, 60(1):3, 2013.

[Ioa07] Lawrence M Ioannou. Computational complexity of the quantum separability problem. *Quantum Information & Computation*, 7(4):335–370, 2007.

[IP99] Russell Impagliazzo and Ramamohan Paturi. Complexity of k-sat. In *Computational Complexity, 1999. Proceedings. Fourteenth Annual IEEE Conference on*, pages 237–240. IEEE, 1999.

[JJUW10] Rahul Jain, Zhengfeng Ji, Sarvagya Upadhyay, and John Watrous. Qip= pspace. *Communications of the ACM*, 53(12):102–109, 2010.

[Kho02] Subhash Khot. On the power of unique 2-prover 1-round games. In *Proceedings of the thiry-fourth annual ACM symposium on Theory of computing*, pages 767–775. ACM, 2002.

[KKR06] Julia Kempe, Alexei Kitaev, and Oded Regev. The complexity of the local hamiltonian problem. *SIAM Journal on Computing*, 35(5):1070–1097, 2006.

[KM02] Hirotada Kobayashi and Keiji Matsumoto. Quantum multi-prover interactive proof systems with limited prior entanglement. In *Algorithms and Computation*, pages 115–127. Springer, 2002.

[KMY01] Hirotada Kobayashi, Keiji Matsumoto, and Tomoyuki Yamakami. Quantum certificate verification: Single versus multiple quantum certificates. *arXiv preprint quant-ph/0110006*, 2001.

[KMY03] Hirotada Kobayashi, Keiji Matsumoto, and Tomoyuki Yamakami. Quantum merlin-arthur proof systems: Are multiple merlins more helpful to arthur? In *Algorithms and Computation*, pages 189–198. Springer, 2003.

[KSV02] Alexei Yu Kitaev, Alexander Shen, and Mikhail N Vyalyi. *Classical and quantum computation*, volume 47. American Mathematical Society Providence, 2002.

[LCV07] Yi-Kai Liu, Matthias Christandl, and Frank Verstraete. Quantum computational complexity of the N-representability problem: QMA complete. *Physical review letters*, 98(11):110503, 2007.

[Lev73] Leonid A Levin. Universal sequential search problems. *Problemy Peredachi Informatsii*, 9(3):115–116, 1973.

[LGNN11] Francois Le Gall, Shota Nakagawa, and Harumichi Nishimura. On QMA protocols with two short quantum proofs. *arXiv preprint arXiv:1108.4306*, 34:38, 2011.

[MW05] C. Marriott and J. Watrous. Quantum Arthur–Merlin games. *Computational Complexity*, 14(2):122–152, 2005.

[NC11] M. Nielsen and I. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, Cambridge, 10th anniversary edition edition, 2011.

[OT08] Roberto Oliveira and Barbara M. Terhal. The complexity of quantum spin systems on a two-dimensional square lattice. *Quantum Info. Comput.*, 8(10):900–924, November 2008.

[Per12] Attila Pereszlényi. Multi-prover quantum merlin-arthur proof systems with small gap. *arXiv preprint arXiv:1205.2761*, 2012.

[Sha92] Adi Shamir. Ip= pspace. *Journal of the ACM (JACM)*, 39(4):869–877, 1992.

[SW15] Yaoyun Shi and Xiaodi Wu. Epsilon-net method for optimizations over separable states. *Theoretical Computer Science*, 2015.

[Wat99] John Watrous. Pspace has constant-round quantum interactive proof systems. In *Foundations of Computer Science, 1999. 40th Annual Symposium on*, pages 112–119. IEEE, 1999.

[Wat00] John Watrous. Succinct quantum proofs for properties of finite groups. In *Foundations of Computer Science, 2000. Proceedings. 41st Annual Symposium on*, pages 537–546. IEEE, 2000.